REPORT DOCUMENTATION PAGE

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE September 2000	3. REPORT TYPE AND DATES COVERED Final Report

4. TITLE AND SUBTITLE

Asynchronous Systems Architecture Project

5. FUNDING NUMBERS

DAAH04-94-G-0274

6. AUTHOR(S)
Alain J. Martin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
California Institute of Technology
Pasadena, CA 91125

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 33615.4-EL

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT

The main result of this research has been to design a new method for high preformance asynchronous circuits based on very fine pipelining. The method has been demonstrated succesfully through the design of two main chips: an asynchronous lattice-structure digital filter and the MiniMIPS.

With 2 million transistors, the asynchronous MiniMIPS is probably the largest VLSI chip and the fastest microprocessor successfully designed in academia. It was found to be entirely functional on first silicon. It is about two-and-a-half the performance of commercial microprocessors of the same type and in equivalent technologies. Performance in excess of 400 MIPS are expected in TSMC's 0.25μ CMOS technology, which is remarkable given that the processor issues instructions one at a time while all processors in this performance range use multiple instructions issue

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

FINAL PROGRESS REPORT
to the
ARMY RESEARCH OFFICE

Asynchronous Systems Architecture Project
Alain J. Martin
California Institute of Technology
Grant DAAH04-94-G-0274

Foreword

This project fits in a sequence of DARPA-funded projects at Caltech in the area of VLSI design. The performance period for this project was from September 11, 1994 until September 10, 1997. Two no-cost extensions were granted: the first one from September 11, 1997 until September 10, 1998; the second one from September 11, 1998 until March 10, 1999.

20001124 003

Contents

Objective

The objective of this project has been to develop a complete design methodology for asynchronous VLSI systems, including basic circuit theory, synthesis method, optimization techniques, logical and electrical simulations, CAD tools, testing, microprocessor architectures for high performance and low power. The motivation for this long-term effort is that, in deep submicron technology, asynchronous circuits offer important advantages compared to synchronous design, in terms of logical and electrical robustness, performance, power consumption, and most importantly design efficiency. These advantages are even more critical in emerging and future technologies.

Asynchronous techniques are based on local communication among concurrent units. Communication and synchronization among units are imple-

mented by handshake protocols. There is no concept of global time—no clocks are used—and no assumptions are made about the duration of an action (delay-insensitivity). Several advantages derive from this approach.

1. All design problems associated with clocks are eliminated.

2. DI circuits are very robust to variations in physical parameters caused by fabrication, scaling, modeling errors, metastability, or variations in temperature or voltage.

3. Asynchronous techniques have a low-power advantage due to the absence of clocks, the locality of activity, the automatic shut-off of inactive parts, and the absence of spurious transitions.

4. The modularity, scalability, and weak synchronization constraints of asynchronous techniques drastically simplify the overall design of large systems.

The specific objective of this project was to demonstrate the advantages of asynchronous techniques for the design of both high-performance and low-power digital VLSI systems, in particular general-purpose microprocessors.

Approach

The main vehicle for this demonstration has been the design of an asynchronous clone of a commercial microprocessor, the MIPS R3000. A performance advantage in terms of $E * t^2$ (the energy per instruction times the square of the cycle time) compared to synchronous designs is expected, which will demonstrate that an asynchronous design with a good $E * t^2$ can outperform low-power synchronous designs at low voltage and high-performance synchronous designs at high voltage.

A new generation of CAST, a suite of CAD tools for the design of asynchronous circuits has been developed. CAST includes logic synthesis, performance analysis, digital, analog, and mixed simulation, and layout tools.

3

Summary of Results

The main result of this research has been to design a new method for high preformance asynchronous circuits based on very fine pipelining. The method has been demonstrated succesfully through the design of two main chips: an asynchronous lattice-structure digital filter and the MiniMIPS.

With 2 million transistors, the asynchronous MiniMIPS is probably the largest VLSI chip and the fastest microprocessor successfully designed in academia. It was found to be entirely functional on first silicon. It is about two-and-a-half the performance of commercial microprocessors of the same type and in equivalent technologies. Performance in excess of 400 MIPS are expected in TSMC's 0.25μ CMOS technology, which is remarkable given that the processor issues instructions one at a time while all processors in this performance range use multiple instructions issue

The Caltech Filter

The Caltech filter is an asynchronous circuit for a programmable FIR filter of 1 to 80 coefficients. It was designed in 1994 to test the new design ideas before embarking on the design of the MiniMIPS. Samples and coefficients are 12 bits. The chip has 250K transistors in MOSIS's HP 0.9μm technology. Measurements show that the chips work correctly from 1V to 5V and from 77K to 400K. At 1.1V, the chip operates at 36 Mops and consumes 0.02W. Let us compare it to the low-power Pleiades chip from Berkeley. Pleiades was fabricated in 0.5μm CMOS. According to the information on the web page, the performances are 3mW at 14 Mops with a word size of 16 bits.

The $E * t^2$ of the Caltech filter is $0.4 * 10^{-24}$. The $E * t^2$ of Pleiades is 10^{-24}.

Adjusted for feature size and word size, the performance of the Caltech filter is more than a factor 10 better than that of Pleiades, even though the Caltech filter was not designed for low-power.

An Asynchronous MIPS R3000 Microprocessor

The asynchronous MIPS R3000 microprocessor (the "MiniMIPS") project had two main goals: first, investigating issues in asynchronous processor architecture that had not been tackled yet (caches, precise exceptions, register

4

bypassing, branch-delay slot, and branch prediction), secondly, developing new techniques for high-performance asynchronous digital VLSI.

The first prototype was fabricated in a 0.6 μm SCMOS process (Hewlett-Packard via MOSIS). Except for the TLB and partial-word instructions, this prototype implements the complete ISA specification of the MIPS. With 2 million transistors, this chip is probably the largest VLSI chip and the fastest microprocessor successfully designed in academia. It was found to be entirely functional on first silicon.

According to extensive SPICE simulation, at 3.3 V and 75 °C, the Min-iMIPS was expected to dissipate 8 W and run at 280 MIPS. At 2.0 V and 75 °C, it should dissipate 1 W and run at 150 MIPS. It should exceed 300 MIPS at 25 °C. The transistor count is 1.25M for the caches, and 750K for the datapath and control. The first prototype doesn't meet the expected performance: This specific HP run was 20% slower than normal, and a layout error left a long poly wire in a critical path slowing down the critical path by another 20%.

The prototype runs at 180 MIPS and consumes 4W at 3.3 V and 25 °C. It is about two-and-a-half times the performance of commercial microprocessors of the same type and in equivalent technologies. Performance in excess of 400 MIPS are expected in TSMC's 0.25μm CMOS technology, which is remarkable given that the processor issues instructions one at a time while all processors in this performance range use multiple instructions issue

Comparing the performance of the MiniMIPS to that of existing synchronous and asynchronous processors is not straightforward. High-performance versions of the MIPS R3000 do not exist, and high-performance processors fabricated in similar technologies are quite different. Furthermore, the technologies used in current microprocessor designs are not always directly comparable to the 0.6 μm SCMOS process (Hewlett-Packard via MOSIS) used for the MiniMIPS. A comparison of the MiniMIPS to the Orion R4600, the DEC StrongARM, and the AMULET2 is summarized in the table below. By MIPS, we mean millions of R3000 or R4000 instructions per second for the MiniMIPS and the Orion. The numbers for the AMULET and StrongARM refer to their published Dhrystone MIPS figures.

The graph of Figure 1 offers another way to compare performances. The solid line is the Iso-$E * t^2$ line of the MiniMIPS in 0.6 μm. The dotted line is an adjustment of the MiniMIPS Iso-$E * t^2$ line for 0.35 μm technology. The points represent data for different processors. The scale is a log-log
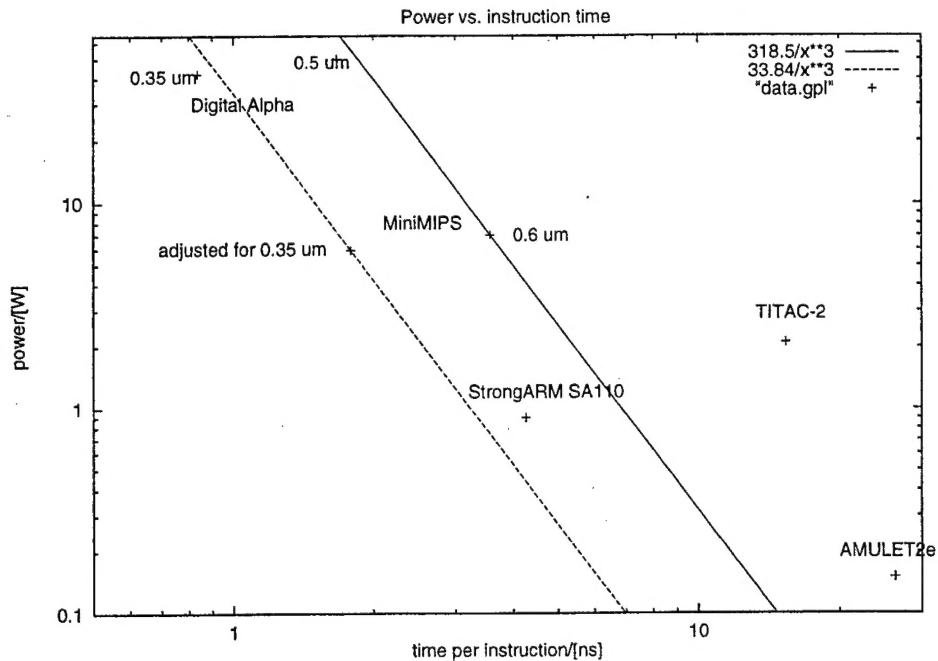
Figure 1: Comparison of $E * t^2$ figures for different microprocessors. The solid ligne represents the iso-$E * t^2$ ligne for the MiniMIPS in loglog scale. The dotted ligne is the iso-$E * t^2$ adjusted for 0.35 micron technology. The points closer to the origin correspond to better performance.

scale, hence the straight lines. We have used the metric $E * t^2$ because it is independent of voltage, to first order, and therefore permits us to compare designs that operate at differing voltages. Titac2 is an asynchronous MIPS microprocessor designed at the University of Tokyo.

High-Performance Asynchronous Techniques

The performance targeted for the asynchronous MIPS required a drastic improvement of the asynchronous pipeline mechanism. A new ultra-fine pipeline mechanism was introduced whose high throughput and low forward latency bring the performance of asynchronous circuits on a par with the best clocked dynamic pipelines.

In an asynchronous system, since there is no clock signal, a pipeline stage

or functional unit needs to generate its own "completion signal" to indicate the completion of that stage and the availability of the result for the next stage. The generation of a completion signal takes a time proportional to $logN$, where N is the number of bits of output data. Since a critical cycle usually contains 2 completion signals, for 32 or 64 bits of output, the performance of a pipeline including such a mechanism is seriously limited. The completion detection problem is the cause of the remaining skepticism against asynchronous techniques.

This problem was solved by pipelining the completion mechanism. The data path is decomposed in small units - say, 4 bit wide - that generate a completion signal through a small, constant, delay. The collection of all the unit completion signals is pipelined, and they do not appear on any critical cycle. Now, the completion detection overhead is reduced to a small constant. By increasing the pipelining, this method could increase the latency on pipeline stalls, and therefore requires additional care in the design of the datapath, and the reduction of the forward latency.

In the current design style, the forward latency of a pipeline stage is only 1/6 of the cycle time, and the cycle time is decreased by a combination of deep pipelining and pipelined completion. The frequencies achieved with this approach are very difficult to achieve for clocked designs due to clock frequency limitations. For example, the latency of a bare pipeline stage in 0.6μm CMOS is 500 picosecs, which corresponds to a 2 GigaHertz clock frequency. Similarly, the low forward latencies we now achieved can be difficult to match in clocked design because of the margins required for proper clock operation.

Other Results

We developed a new optimization technique, called "slack matching", to adjust the amount of buffering present in the channel. In a finely pipelined asynchronous design like the MIPS, the overall throughput performance is critically dependent on the amount of slack in the different linear and circular pipelines that compose the architecture. Adjusting the slack of pipelines for performance optimization is what we call slack matching.

We also devised necessary and sufficient conditions under which the slack in an asynchronous circuit can be modified without affecting correctness. The results let the designer reason about a complex pipelined processor by

7

treating it as an unpipelined system. The results were used to demonstrate the correctness of new program transformations for slack adjustment used in the design of the MIPS.

A novel processor architecture, called the Branch Processor, was invented to address the problem of fetching instructions from memory. The architecture includes a separate stream of instructions for controlling program counter computation. It allows performance gains from being able to fetch more instructions at once, reduced code size, and increased cache hit rates.

We designed a new synchronizer which can be used to sample the value of a wire in an asynchronous system. The circuit functions correctly and is guaranteed to produce an output as long as the difference between Vdd and GND is more than the threshold of the switching devices.

Finally, a systematic approach to eliminating charge sharing in asynchronous circuits was developed. Because a signal in an asynchrnous circuit must change monotonically from one voltage level to the other, eliminating charge sharing is critical. (The report on this result has not yet been written.)

CAD Tools

A new suite of CAD tools was developed to speed up the development of high-performance asynchronous circuits. At the electrical simulation level, a new analog simulator has been developed. In conjunction with the new analog simulator, tools have been developed to examine automatically the simulation traces and to pinpoint circuit problems with minimal human intervention. These software tools are also able to speed up circuit performance tuning.

We designed a new production-rule simulator. (Productio rules are the last level of logical design in our compilation method.) This new simulator is currently being used to create a combined digital-analog simulator.

We also designed a verification tool that can be used to check the correctness of physical layout against the last level of logical design (production rules) in our compilation method. The program is tailored for asynchronous circuits, but can also be used (and has been used) to check synchronous circuits.

8

List of Publications and Technical Reports

1. **Solving Certain Large Instances of the Quadratic Problem: Steinberg's Examples**
   Mika Nyström, Caltech Technical Report CS-TR-99-06.

2. **A Fail-Safe Synchronizer**
   with Mika Nyström and Rajit Manohar, submitted for publication, August 2000.

3. **A Pipelined Asynchronous Cache System**
   with Mika Nyström and Andrew Lines, submitted for publication, August 2000.

4. **An Asynchronous Approach to Energy-efficient Computing and Communication**
   in *Proceedings of SSGRR 2000, International Conference on Advances in Infrastructure for Electronic Business, Science, and Education on the Internet*, L'Aquila, Italy, August 2000.

5. **The Entropy of Traces in Parallel Computation**
   Rajit Manohar, in *IEEE Transactions on Information Theory*, 45(5), July 1999.

6. **Asynchronous Parallel Prefix Computation**
   Rajit Manohar and José A. Tierno, in *IEEE Transactions on Computers, 47(11):1244–1252*, November 1998.

7. **Projection: A Synthesis Technique for Concurrent Systems**
   with Rajit Manohar and Tak-Kwan Lee. in *Proceedings of the Fifth International Symposium on Advanced Research in Asynchronous Circuits and Systems*, April 1999.

8. **The Design of an Asynchronous MIPS R3000**
   Alain J. Martin, Andrew Lines, Rajit Manohar, Mika Nyström, Paul Penzes, Robert Southworth, Uri V. Cummings, and Tak-Kwan Lee. *Proceedings of the 17th Conference on Advanced Research in VLSI*, pp. 164–181, September 1997.

9. **Slack Elasticity in Concurrent Computing**
   with Rajit Manohar, in *Proceedings of the Fourth International Conference on the Mathematics of Program Construction*, Lecture Notes in Computer Science 1422, pp. 272-285, Springer-Verlag 1998.

10. **A Program Transformation Approach to Asynchronous VLSI Design**
    in *NATO ASI Series, Deductive Program Design*, Broy Ed., Springer Verlag, (1996)

11. **The Energy and Entropy of VLSI Computations**
    with Jose A. Tierno and Rajit Manohar, in *Proceedings of the Second International Symposium on Advanced Research in Asynchronous Circuits and Systems*. March 1996.

12. **Quasi-delay-insensitive circuits are Turing-complete**
    Rajit Manohar and Alain J. Martin, Invited article, *Second International Symposium on Advanced Research in Asynchronous Circuits and Systems. March 1996*. Available as Caltech technical report CS-TR-95-11, November 1995.

13. **Asynchronous VLSI Design**
    Tutorial, *1995 Israel Workshop on Asynchronous VLSI*, March 1995.

14. **An Action System Specification of the Caltech Asynchronous Microprocessor**
    with R.J.R. Back and K.Sere, in *Proceedings of 3rd International Conference on the Mathematics of Program Construction*, Lecture Notes in Computer Science LCNS947, Springer-Verlag, 1995.
    to be republished in *Science of Computer Programming*

15. **Extended Event-rule Systems and Performance Analysis of Asynchronous Systems**
    with Tony Lee, in *Proceedings of TAU95, ACM International workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, November 1995.

16. **Low-Energy Asynchronous Memory Design**
    with Jose A. Tierno, in *Proceedings of International Symposium on*

Advanced Research in Asynchronous Circuits and Systems, IEEE Computer Society Press, 176-185, 1994.

17. **An Asynchronous Pipeline Lattice-structure Filter**
with Uri Cummings and Andrew Lines, in *Proceedings of International Symposium on Advanced Research in Asynchronous Circuits and Systems*, IEEE Computer Society Press, 126-133, 1994.

18. **Robustness Issues in Asynchronous VLSI Techniques**
in *International Symposium on Fault-Tolerant Computing (FTCS-24)* June 1994, Tokyo, Japan volume 94, Nr.128, 1994

PH.D. THESES

1. **A General Approach to Performance Analysis and Optimization of Asynchronous Circuits**, Lee, Tak Kwan, 18 May 1995.

2. **An Energy Complexity Model for VLSI Computations**, Tierno, José Andres, 11 January 1995.

3. **Semantics for VLSI Synthesis**, van der Goot, Marcel René, 26 May 1995.

4. **The Impact of Asynchrony on Computer Architecture**, Rajit Manohar, 25 June 1998.

MASTER'S THESES

1. **The CAST Parser**, Matthew S. Hanna, 18 April 2000.

2. **Gretel: An Interactive Router for Magic**, Eitan Eduardo Grinspun, October 1, 1999.

3. **Pipelined Asynchronous Cache Design**, Mika Nystrom, April 1999.

4. **Efficient Compilation of Handshaking Programs to Production Rules**, Robert Southworth, May 2000.

5. **A Graphical Method for Process Decomposition**, Catherine Grace Wong, May 2000.

6. **Pipelined Asynchronous Circuits**, Andrew Matthew Lines, June 1998.

7. **The Design of High-performance Asynchronous Circuits for the Caltech MiniMIPS Processor**, Paul Pénzes, December 1998.

8. **On Detection and Generation of Deadlock-free Reshufflings in VLSI Synthesis Method**, Jiazhao Jessie Xu, May 26, 1995.

List of Personnel

1. Tak Kwan Lee, Ph.D., May 1995.

2. José Andres Tierno, Ph.D., January 1995.

3. Marcel René van der Goot, Ph.D., May 1995.

4. Rajit Manohar, Ph.D., June 1998.

5. Matthew S. Hanna, MS, April 2000.

6. Eitan Eduardo Grinspun, MS, October 1999.

7. Mika Nystrom, MS, April 1999.

8. Robert Southworth, MS, May 2000.

9. Catherine Grace Wong, MS, May 2000.

10. Andrew Matthew Lines, MS, June 1998.

11. Paul Pénzes, MS, December 1998.

12. Jiazhao Jessie Xu, MS, May 1995.

Report of Inventions

Patents

1. Rajit Manohar and Alain J. Martin. Parallel Prefix Operations in Asynchronous Processors. US Patent No. 5,999,961, December 1999.

2. Uri V. Cummings, Andrew Lines, Alain J. Martin. Pipelined Completion for Asynchronous Communication. US Patent No. 6,038,656.

Patents Pending

1. CIT No. 2654 Application Type: Non Provisional Title: Improved Devices and Methods for Asynchronous Processing Inventors: Uri V. Cummings; Tak K. Lee; Andrew Lines; Rajit Manohar; Alain J. Martin; Mika Nystrom; Paul Penzes; Robert Southworth. Date Filed: 7/16/98 Serial No. 09/118,140.

2. CIT No. 2654-F Application Type: PCT Title: The Design of an Asynchronous MIPS R3000 Microprocessor Inventors: Uri V. Cummings; Tak K. Lee; Andrew Lines; Rajit Manohar; Alain J. Martin; Mika Nystrom; Paul Penzes; Robert Southworth. Date Filed: 7/16/98 Serial No. PCT/US98/14843

3. CIT No. 2654-P Application Type: Provisional Title: The Design of an Asynchronous MIPS R3000 Microprocessor Inventors: Uri V. Cummings; Tak K. Lee; Andrew Lines; Rajit Manohar; Alain J. Martin; Mika Nystrom; Paul Penzes; Robert Southworth. Date Filed 7/16/97 Serial No. 60/052,714

4. CIT No. 2709 Application Type: Non-Provisional Title: Exception Processing in Asynchronous Processor. Inventors: Rajit Manohar; Alain J. Martin. Date Filed 9/15/98 Serial No. 09/153,756

5. CIT No. 2709P Application Type: Provisional Title: Exception Processing in Asynchronous Processor. Inventors: Rajit Manohar; Alain J. Martin. Date Filed 9/15/97 Serial No. 60/058,995

6. CIT No. 2846 Application Type: Non Provisional Title: Reshuffled Communications Processes in Pipelined Asynchronous Circuits. Inventors: Andrew M. Lines; Alain J. Martin Date Filed: 7/22/99 Serial No. 09/360,468

7. CIT No. 2846-C1 Application Type: CONT Title: Reshuffled Communications Processes in Pipelined Asynchronous Circuits. Inventors: Andrew M. Lines; Alain J. Martin Date Filed: 2/10/00 Serial No. 09/501,638.

8. CIT No. 2846-P Application Type: Provisional Title: Pipelined Asynchronous Circuits. Inventors: Andrew M. Lines; Alain J. Martin Date Filed; 7/22/98 Serial No. 60/093,840.

9. CIT No. 2846-PCT Application Type: PCT Title: Reshuffled Communications Processes in Pipelined Asynchronous Circuits. Inventors: Andrew M. Lines; Alain J. Martin Date Filed: 7/22/98 Serial No. PCT/US99/16628

10. CIT No. 2856 Application Type: Non Provisional Title: Branch Processor for Architecture. Inventors: Rajit Manohar; Alain J. Martin. Date filed 8/20/99 Serial No. 09/378,330

11. CIT No. 2856-P Application Type: Provisional Title: Processing Element with Special Application for Branch Functions. Inventors: Rajit Manohar; Alain J. Martin. Date filed 8/21/98 Serial No. 60/097,515.

12. CIT No. 2856-PCT Application Type: PCT Title: Processing Element with Special Application for Branch Functions. Inventors: Rajit Manohar; Alain J. Martin. Date filed 8/20/99 Serial No. PCT/US99/19197